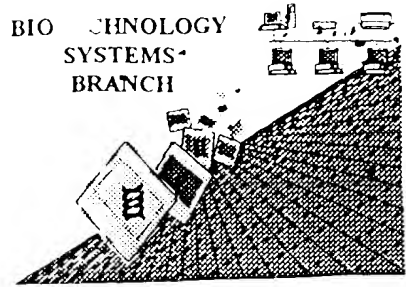


2570  
0803

BIO TECHNOLOGY  
SYSTEMS  
BRANCH



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/855294A

Source: OIE

Date Processed by STIC: 08/08/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

### Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST 25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 09/055294A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos  
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2      Invalid Line Length  
The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering  
The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII  
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length  
Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"  
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)  
Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped  
  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8      Skipped Sequences  
    (NEW RULES)  
Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
<210> sequence id number  
<400> sequence id number  
000
- 9 ✓ Use of n's or Xaa's  
    (NEW RULES)  
Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10      Invalid <213>  
    Response  
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11      Use of <220>  
Sequence(s)      missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"  
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n  
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH - Biotechnology Systems Branch - 08/21/2001

The type of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors

OIPE

## RAW SEQUENCE LISTING

DATE: 08/08/2001

PATENT APPLICATION: US/09/855,294A

TIME: 10:48:05

Input Set : A:\ES.txt

Output Set: N:\CRF3\08082001\I855294A.raw

4 <110> APPLICANT: Croce, Carlo  
 5 Brenner, Charles  
 6 Pekarski, Yuri  
 10 <120> TITLE OF INVENTION: CRYSTAL STRUCTURE OF WORM NitFhit  
 11 REVEALS THAT A Nit TETRAMER BINDS TWO Fhit DIMERS  
 14 <130> FILE REFERENCE: CRO01.NP007  
 16 <140> CURRENT APPLICATION NUMBER: 09/855,294A  
 17 <141> CURRENT FILING DATE: 2001-05-15  
 19 <150> PRIOR APPLICATION NUMBER: 60/204,713  
 20 <151> PRIOR FILING DATE: 2000-05-16  
 22 <160> NUMBER OF SEQ ID NOS: 11  
 24 <170> SOFTWARE: FastSEQ for Windows Version 4.0

## ERRORED SEQUENCES

346 <210> SEQ ID NO: 8  
 347 <211> LENGTH: 1359  
 348 <212> TYPE: DNA  
 349 <213> ORGANISM: Homo sapien  
 351 <400> SEQUENCE: 8

352	gtggtgcttg	tctgcagagt	catgaacctct	ttccgcttgg	ccctcatcca	gcttcagatt	60
353	tcttccatca	aatcagataa	cgtcaactgc	gcttgtagct	tcattccggga	ggcagcaacg	120
354	caaggagcca	aaatagtttc	tttgccggaa	tgctttaatt	ctccatatgg	agcgaaatat	180
355	tttccctgaat	atgcagagaa	aattcctggg	gaatccacac	agaagctttc	tgaagtagca	240
356	aaggaatgca	gcataatatct	cattggaggc	tctatccctg	aagaggatgc	tgggaaatta	300
357	tataacacct	gtgctgtggt	tgggcctgat	ggaactttac	tagcaaagta	tagaaagatc	360
358	catctgtttg	acattgatgt	tcctggaaaa	attacatttc	aagaatctaa	aacattgagt	420
359	ccgggtgata	gttttccac	atttgatact	ccttactgca	gagtgggtct	gggcatctgc	480
360	tacgacatgc	ggtttgcaga	gcttgcacaa	atctacgcac	agagaggctg	ccagctgttg	540
361	gtatatccag	gagcttttaa	tctgaccact	ggaccagccc	attgggagtt	acttcagcga	600
362	agccgggctg	ttgataatca	ggtgtatgtg	gccacagcct	ctcctgcccg	ggatgacaaa	660
363	gcctcctatg	ttgcctgggg	acacagcacc	gtggtgaacc	cttgggggga	ggttctagoc	720
364	aaagctggca	cagaagaagc	aatcgtgtat	tcagacatag	acctgaagaa	gctggctgaa	780
365	atacgccagc	aaatccccgt	ttttagacag	aagcgatcag	acctctatgc	tgtggagatg	840
366	aaaaagccct	aaagtttatg	tttctaattg	gtcacagaat	aggacgatat	gattctacaa	900
367	cataatcaac	tcctatttaa	attctttaat	gaagaaaaaa	aatttaaaaa	aaaaaaaaaa	960
368	aacctagggt	ctctatttag	atgagaaagc	ctcattatgc	tgacattttc	cacgccacat	1020
369	taaatagtta	aaaaggatgc	agcctggagc	cagagagcag	aaagctgggc	tggttctgaa	1080
370	gcttcttcca	tacttaagtt	gcctccaagc	agtttgtgaa	agtatcagat	cttgggtatcc	1140
371	tgggtgattga	ttcacctaata	ataaatatat	ttgtgtcatg	aacctcttaa	aaagttgctg	1200
372	ggagttgtaa	tctccatcat	ctaggaaaac	gtgggtctgg	gtgctattct	tttccaagca	1260
E--> 373	ggtaccttga	aagttccatt	tttgggttca	tgagttagcta	taggaacgca	agggtgatac	1320
374	atctttgggt	gttttgccag	agaagttggg	cagccccac			1359

*Errored Sequence Listing: You must enumerate  
 unknown nucleotides in field 221, 222 and 223*

*Errored*

The type of errors shown exist throughout  
 the Sequence Listing. Please check subsequent  
 sequences for similar errors.

## RAW SEQUENCE LISTING

DATE: 08/08/2001

PATENT APPLICATION: US/09/855,294A

TIME: 10:48:05

Input Set : A:\ES.txt

Output Set: N:\CRF3\08082001\I855294A.raw

408 &lt;213&gt; ORGANISM: X. laevis

410 &lt;400&gt; SEQUENCE: 10

```

411 gtgagtgacg tgtgcgcagt ggcaactaag gcctcctggg aaaatgtaga ggagcgtgag      60
412 cttcgcggga caggacaggg tcttaggctc tgccttggtt ccacacgccc ttgtgcagac      120
413 tgctatagac tgtgacttta accctgtgtc cggatatagg ggttagaagc ctgagtgcaa      180
414 tggtcgtgtc ccacaagccc ctgattgccg tgtgccagat gacttcaacc tctgataagg      240
415 agaagaattt cgccacgtgt tcgcggtga tccgggaggc tgcggggcgt cgcgcttgca      300
416 tgggtgtttct gccggaagcc tttgactata tcgggggcag cattgaggag acgctgagtc      360
417 tggctgagtc tctacatggg gacaccattc agcgttacac ccaactcgcc agggagtgtg      420
418 ggctctggct ttccctgggg ggatttcatg agaaaggacc caactgggac acggaccaac      480
419 gcatttccaa ttctcacgtg gttgtggaca acacagggca catagtatcg gtgtaccgca      540
420 aggtcacctt gtttgacgta gacttgaga atggagtgtc actcagagag agcagttcca      600
421 cctcccccgg agcagagctt attcgcccca tcaattctcc agcaggaaaag attggcctgg      660
422 ggggtgtgta cgacctccgc ttcccagaat tctccttggc tctggcccaa caaggagcag      720
423 aacttctcac ttacccttct gccttcaccc tcaactactg tctggcacat tgggaggtgt      780
424 tgctgagagc ccgtgccata gaaacccagt gctacgtagt tgacgcggca cagacagaca      840
425 gacacaatga gaagaggacg tcctatggtc acgctatggt ggtagaccgg tgggggctgg      900
426 tcattggcca atgccaggaa ggaacaggaa tatgttatgc tgagattgac attccctaca      960
427 tggagcgtgt gaggcgggac atgcgggtgt ggaggcaccg caggactgat ctgtatggga     1020
428 aaatctcctt taataaaccg gactgactcc ataatggatc acctgcacct atggggggcaa     1080
E--> 429 agnctttccc ctgattgctg aaattcctca atctgtgact gtgaatgaca atgaacgtga     1140
430 cttggaattg cctgggttatg gcaccggcaa tgattctcta cagtaattct caataaaagt     1200
431 gctgaacctt aaaa                                     1214

```

Errors

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/855,294A

DATE: 08/08/2001

TIME: 10:48:06

Input Set : A:\ES.txt

Output Set: N:\CRF3\08082001\I855294A.raw

L:373 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:8

L:429 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:10

**STATISTICS SUMMARY**

PATENT APPLICATION: US/09/855,294A

DATE: 08/08/2001

TIME: 10:48:06

Input Set : A:\ES.txt

Output Set: N:\CRF3\08082001\I855294A.raw

Application Serial Number: US/09/855,294A

Alpha or Numeric: Numeric

Application Class:

Application File Date: 05-15-2001

Art Unit: OIPE

Software Application: FastSeq

Total Number of Sequences: 11

Total Nucleotides: 3865

Total Amino Acids: 2382

Number of Errors: 2

Number of Warnings: 0

Number of Corrections: 0

**MESSAGE SUMMARY**

340 W: 2 ((46) "n" or "Xaa" used: Feature required)